

CRFI #12/11/00

ENTERED

See page 6



1600

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/09/865,018A

TIME: 07:04:56

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt

Output Set: N:\CRF4\02142003\I865018A.raw

4 <110> APPLICANT: Massague et al.
 5 <120> TITLE OF INVENTION: ISOLATED p27 PROTEIN AND METHODS FOR ITS PRODUCTION AND USE
 8 <130> FILE REFERENCE: GPCI-P08-079
 10 <140> CURRENT APPLICATION NUMBER: 09/865018A
 11 <141> CURRENT FILING DATE: 2001-05-24
 12 <150> PRIOR APPLICATION NUMBER: 08/854039
 14 <151> PRIOR FILING DATE: 1997-05-09
 16 <160> NUMBER OF SEQ ID NOS: 12
 18 <170> SOFTWARE: PatentIn version 3.1
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 534
 22 <212> TYPE: DNA
 23 <213> ORGANISM: Histella vison
 25 <220> FEATURE:
 26 <221> NAME/KEY: CDS
 27 <222> LOCATION: 1..(534)
 28 <223> OTHER INFORMATION:

RECEIVED

FEB 19 2003

TECH CENTER 1600/2900

W--> 30 <400> 1

31 atg tca aac gtg cgg gtg tct aac ggg agc ccg agc ctg gag cgg atg	48
32 Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met	
33 1 10 15	
35 gac gcc aga cag cgg gag tac ccc aag ccc tcc gcc tgc aga aac ctc	96
36 Asp Ala Arg Gln Ala Glu Tyr Pro Lys Pro Ser Ala Cys Arg Asn Leu	
37 20 25 30	
39 ttc cgc ccc gtc aac cac gaa gag ctg acc cgg gac ttg gag aag cac	144
40 Phe Gly Pro Val Asn His Glu Glu Leu Thr Arg Asp Leu Glu Lys His	
41 35 40 45	
43 cgc aga gac atg gaa gag gca agc cag cgc aag tgg aat ttt gat ttc	192
44 Arg Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe	
45 50 55 60	
47 cag aat cac aag ccc ctg gag gcc aca tac gag tgg cag gag ttg gag	240
48 Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu	
49 65 70 75 80	
51 aag ggc agc ttg cgg gag ttc tac tac aga ccc cgg cgg cca ccc aaa	288
52 Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys	
53 85 90 95	
55 ggc gcc tcc aag gtg cgg cgg cag gag agc cag gac gtc agc ggg acc	336
56 Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Thr	
57 100 105 110	
59 cgg cag gcc ttc ttt tta atg ggt tct cag gcc aac ttc gag gac aca	384
60 Arg Gln Ala Val Pro Leu Met Gly Ser Gln Ala Asn Ser Glu Asp Thr	
61 115 120 125	
63 cac ttt gta gac gaa aac atc gac aag ggt gcc aac cag gct ggc tta	432

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/09/865,018A

TIME: 07:04:56

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt

Output Set: N:\CRF4\02142003\I865018A.raw

```

64 His Leu Val Asp Gln Lys Thr Asp Thr Ala Asp Asn Gln Ala Gly Leu
65      130      135      140
66 ggc gag cag tcc act ggg atc agg aag cga ccc gcc aca gac gat tcc      480
67 Ala Glu Gln Cys Thr Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
68 115      150      155      160
69 tct cct caa aac aaa aga gcc aac aga aca gaa gaa aat gtc tca gac      528
70 Ser Pro Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
71      165      170      175
72 ggt tcc      534
73 Gly Ser

```

74 <210> SEQ ID NO: 1

80 <211> LENGTH: 178

81 <212> TYPE: PRT

82 <213> ORGANISM: Mustela vison

83 <400> SEQUENCE: 2

```

84 Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
85 1      5      10      15
86 Asp Ala Arg Gln Ala Glu Tyr Pro Lys Pro Ser Ala Cys Arg Asn Leu
87 21      26      31      36
88 Phe Gly Pro Val Asn His Glu Glu Leu Thr Arg Asp Leu Glu Lys His
89 37      42      47      52
90 Arg Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
91 53      58      63      68
92 Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
93 69      74      79      84
94 Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
95 85      90      95      100
96 Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Thr
97 101      106      111      116
98 Arg Gln Ala Val Pro Leu Met Gly Ser Gln Ala Asn Ser Glu Asp Thr
99 117      122      127      132
100 His Leu Val Asp Gln Lys Thr Asp Thr Ala Asp Asn Gln Ala Gly Leu
101 133      138      143      148
102 Ala Glu Gln Cys Thr Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
103 149      154      159      164
104 Ser Pro Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
105 165      170      175

```

106 Gly Ser

107 <210> SEQ ID NO: 3

108 <211> LENGTH: 594

109 <212> TYPE: DNA

110 <213> ORGANISM: Mus musculus

111 <220> FEATURE:

112 <221> NAME/KEY: CDS

113 <222> LOCATION: (1)..(594)

114 <223> OTHER INFORMATION:

W--> 143 <400> 3

```

144 atc tca aac ggc aga ggc tct aac ggc agc ccc agc ctg gag cgg atc      48
145 Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met

```

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/09/865,018A

TIME: 07:04:56

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt

Output Set: N:\CRF4\02142003\I865018A.raw

```

146 1          5          10          15
147 gac gcc aga gaa gcc gat gac ccc aag cct tcc gcc tgc aga aat ctc          96
148 Asp Ala Arg Gln Ala Asp His Pro Lys Pro Ser Ala Cys Arg Asn Leu
149          20          25          30
150 ttc gcc cgg gtc aat cat gaa gaa cta acc cgg gac ttg gag aag cac          144
151 Phe Gly Pro Val Asn His Gln Gln Leu Thr Arg Asp Leu Glu Lys His
152          35          40          45
153 tgc cgg aat atg gaa gaa gag agt cag cgc aag tgg aat ttc gac ttt          192
154 Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
155          50          55          60
156 cag aat cat aag ccc atg gag gcc aga tac gaa tgg cag gag gtc gag          240
157 Gln Asn His Lys Pro Leu Glu Gly Arg Tyr Glu Trp Gln Gln Val Glu
158          65          70          75          80
159 agg gcc aga ttg ccc gag ttc tac tac agg ccc cgg cgc ccc ccc aag          288
160 Arg Gly Ser Leu Pro Gln Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
161          85          90          95
162 agc gcc tgc aag ttg ttg ggc cag gag aga cag gat gtc agc gag agc          336
163 Ser Ala Cys Lys Val Leu Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
164          100          105          110
165 cgc cag gcc ggc cct tta att ggc tct cag gca aac tct gag gac cgc          384
166 Arg Gln Ala Val Pro Leu Ile Gly Ser Gln Ala Asn Ser Gln Asp Arg
167          115          120          125
168 cat ttg ttg gac caa atg cct gac tgg tca gac aat cag gct ggc tta          432
169 His Leu Val Asp Gln Met Pro Asp Ser Ser Asp Asn Gln Ala Gly Leu
170          130          135          140
171 ggc gag cag ctt cca ggc atg agg aag cga cct gct gca gaa gat tat          480
172 Ala Glu Gln Cys Pro Gly Met Arg Lys Arg Pro Ala Ala Glu Asp Ser
173          145          150          155          160
174 tct tgc gaa aac aaa agg gcc aac aga aca gaa gaa aat gtt tca gac          528
175 Ser Ser Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
176          165          170          175
177 ggt tcc cgc aac gct ggc act atg gag cag acg ccc aag aag ccc ggc          576
178 Gly Ser Pro Asn Ala Gly Thr Val Glu Gln Thr Pro Lys Lys Pro Gly
179          180          185          190
180 att cga agc cag aag taa          594
181 Leu Arg Arg Gln Thr
182          195
183 >10> SEQ ID NO: 4
184 >11> LENGTH: 197
185 >12> TYPE: CDS
186 >13> ORGANISM: Mus musculus
187 >14> SEQUENCE: 4
188 Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
189 1          5          10          15
190 Asp Ala Arg Gln Ala Asp His Pro Lys Pro Ser Ala Cys Arg Asn Leu
191          20          25          30
192 Phe Gly Pro Val Asn His Gln Gln Leu Thr Arg Asp Leu Glu Lys His
193          35          40          45
194 Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe

```

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/09/865,018A

TIME: 07:04:56

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt

Output Set: N:\CRF4\02142003\I865018A.raw

```

216      50      55      60
219 Gln Asn His Lys Pro Leu Glu Gly Arg Tyr Glu Trp Gln Glu Val Glu
220 65      70      75      80
223 Arg Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
224      85      90      95
227 Ser Ala Cys Lys Val Leu Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
228      100      105      110
231 Arg Gln Ala Val Pro Leu Ile Gly Ser Gln Ala Asn Ser Glu Asp Arg
232      115      120      125
235 His Leu Val Asp Gln Met Pro Asp Ser Ser Asp Asn Gln Ala Gly Leu
236      130      135      140
239 Ala Glu Gln Cys Pro Gly Met Arg Lys Arg Pro Ala Ala Glu Asp Ser
240      145      150      155      160
243 Ser Ser Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
244      165      170      175
247 Gly Ser Pro Asn Ala Gly Thr Val Glu Gln Thr Pro Lys Lys Pro Gly
248      180      185      190
251 Leu Arg Arg Gln Thr
252      195

```

254 <210> SEQ ID NO: 5

255 <211> LENGTH: 597

256 <212> TYPE: DNA

257 <213> ORGANISM: Homo sapiens

258 <214> FEATURE:

259 <2.1> NAME/KEY: CDS

260 <2.2> LOCATION: (1)..(597)

261 <2.3> OTHER INFORMATION:

W--> 264 <400> 5

```

265 att tca aac ggg cga gtg tct aac gag agc cct agc ctg gag cgg atg      48
266 Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
267 1      5      10      15
269 gag gcc agc cag ggc gag cac ccc aag ccc tgg gcc tgc agg aac ctg      96
270 Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu
271      20      25      30
273 ttc ggc cgg ggg gac cac gaa gag tta acc cgg gac ttg gag aag cac      144
274 Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His
275      35      40      45
277 ttc aga gac atg gaa gag ggc agc cag cgc aag tgg aat ttc gat ttt      192
278 Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
279      50      55      60
281 cag aat cac aaa ccc cta gag ggc aag tac gag tgg caa gag gtg gag      240
282 Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
283 65      70      75      80
285 aac ggc agc ttg ccc gag ttc tac ttc aga cct cgg cgg ccc ccc aaa      288
286 Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
287      85      90      95
289 agt gcc tgc aag gtg cgg cgg cag cag agc cag gat gtc agt gag acc      336
290 Gly Ala Cys Lys Val Pro Ala Gln Ser Gln Asp Val Ser Gly Ser
291      100      105      110

```

RAW SEQUENCE LISTING

DATE: 02/14/2003

PATENT APPLICATION: US/09/865,018A

TIME: 07:04:56

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt

Output Set: N:\CRF4\02142003\I865018A.raw

```

292 cgc cgc ggc ggc cct tta att ggg gct cgc gct aac tct gag gac acg      334
293 Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr
294      115      120      125
297 cat ttg gtg gac cca aag act gat cgc tgc gac agc cag acg ggg tta      432
298 His Leu Val Asp Pro Lys Thr Asp Pro Ser Asp Ser Gln Thr Gly Leu
299      130      135      140
301 ggc gag caa tgc gca gga ata agg aag cga cct gca acc gac gat tct      480
302 Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
303      145      150      155      160
305 tct act caa aac aag aga gcc aac aag aca gaa gaa aat gtt tca gac      528
306 Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
307      165      170      175
309 ggt tcc cca aat gcc ggt tct gtg gag cag acg ccc aag aag cct agc      576
310 Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly
311      180      185      190
313 ctg aga aga cgt caa acg taa      597
314 Leu Arg Arg Arg Gln Thr
315      195
317 <210> SEQ ID NO: 6
318 <211> LENGTH: 198
319 <212> TYPE: PRT
320 <213> ORGANISM: Homo sapiens
321 <400> SEQUENCE: 6
322 Met Ser Asn Val Arg Val Ser Asn Gly Ser Pro Ser Leu Glu Arg Met
323 1      5      10      15
325 Asp Ala Arg Gln Ala Glu His Pro Lys Pro Ser Ala Cys Arg Asn Leu
326      20      25      30
328 Phe Gly Pro Val Asp His Glu Glu Leu Thr Arg Asp Leu Glu Lys His
329      35      40      45
331 Cys Arg Asp Met Glu Glu Ala Ser Gln Arg Lys Trp Asn Phe Asp Phe
332      50      55      60
334 Gln Asn His Lys Pro Leu Glu Gly Lys Tyr Glu Trp Gln Glu Val Glu
335      65      70      75      80
337 Lys Gly Ser Leu Pro Glu Phe Tyr Tyr Arg Pro Pro Arg Pro Pro Lys
338      85      90      95
340 Gly Ala Cys Lys Val Pro Ala Gln Glu Ser Gln Asp Val Ser Gly Ser
341      100      105      110
343 Arg Pro Ala Ala Pro Leu Ile Gly Ala Pro Ala Asn Ser Glu Asp Thr
344      115      120      125
346 His Leu Val Arg Pro Lys Thr Asp Pro Ser Asp Ser Gln Thr Gly Leu
347      130      135      140
349 Ala Glu Gln Cys Ala Gly Ile Arg Lys Arg Pro Ala Thr Asp Asp Ser
350      145      150      155      160
352 Ser Thr Gln Asn Lys Arg Ala Asn Arg Thr Glu Glu Asn Val Ser Asp
353      165      170      175
355 Gly Ser Pro Asn Ala Gly Ser Val Glu Gln Thr Pro Lys Lys Pro Gly
356      180      185      190
372 Leu Arg Arg Arg Gln Thr
373      195

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/865,018A

DATE: 02/14/2003
TIME: 07:04:57

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt
Output Set: N:\CRF4\02142003\I865018A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:13; N Pos. 3,9
Seq#:14; N Pos. 1,13,16,22
Seq#:15; N Pos. 6,9,12,15
Seq#:16; N Pos. 6,9,12
Seq#:17; N Pos. 1,7,10,13,16
Seq#:18; N Pos. 1,10,13,16
Seq#:19; N Pos. 3
Seq#:20; N Pos. 15
Seq#:22; N Pos. 1

VERIFICATION SUMMARY

DATE: 02/14/2003

PATENT APPLICATION: US/09/865,018A

TIME: 07:04:57

Input Set : N:\efs\09865018\GPCI-P08-079SubstituteSequence.txt

Output Set: N:\CRF4\02142003\I865018A.raw

L:30 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:28
L:143 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:141
L:264 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:262
L:490 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:507 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:0
L:518 M:283 W: Missing Blank Line separator, <220> field identifier
L:522 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0
L:538 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0
L:554 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0
L:570 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:0
L:586 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:602 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:0
L:629 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0